Modeling of uncertainties in biochemical reactions
Ljubiša Mišković, Vassily Hatzimanikatis
Laboratory of Computational Systems Biotechnology,
Ecole Polytechnique Fédérale de Lausanne (EPFL), CH–1015 Lausanne, Switzerland.

Correspondence to:
Vassily Hatzimanikatis
EPFL / SB / ISIC / LCSB
CH H4 625 (Bat. CH)
Station 6
CH-1015 Lausanne, Switzerland
Tel: +41 21 693 98 70; Fax: +41 21 693 98 75;
Email: vassily.hatzimanikatis@epfl.ch
Abstract

Mathematical modeling is an indispensable tool for research and development in biotechnology and bioengineering. The formulation of kinetic models of biochemical networks depends on knowledge of the kinetic properties of the enzymes of the individual reactions. However, kinetic data acquired from experimental observations bring along uncertainties due to various experimental conditions and measurement methods. In this contribution, we propose a novel way to model the uncertainty in the enzyme kinetics and to predict quantitatively the responses of metabolic reactions to the changes in enzyme activities under uncertainty. The proposed methodology accounts explicitly for mechanistic properties of enzymes and physico-chemical and thermodynamic constraints, and is based on formalism from systems theory and metabolic control analysis. We achieve this by observing that kinetic responses of metabolic reactions depend: (i) on the distribution of the enzymes among their free form and all reactive states; (ii) on the equilibrium displacements of the overall reaction and that of the individual enzymatic steps; and (iii) on the net fluxes through the enzyme. Relying on this observation, we develop a novel, efficient Monte Carlo sampling procedure to generate all states within a metabolic reaction that satisfy imposed constrains. Thus we derive the statistics of the expected responses of the metabolic reactions to changes in enzyme levels and activities, in the levels of metabolites, and in the values of the kinetic parameters. We present aspects of the proposed framework through an example of the fundamental three-step reversible enzymatic reaction mechanism. We demonstrate that the equilibrium displacements of the individual enzymatic steps have an important influence on kinetic responses of the enzyme. Furthermore, we derive the conditions that must be satisfied by a reversible three-step enzymatic reaction operating far away from the equilibrium in order to respond to changes in metabolite levels according to the irreversible Michelis-Menten kinetics. The efficient sampling procedure allows easy, scalable, implementation of this methodology to modeling of large-scale biochemical networks.

Introduction

Advancements in genome sequencing have sparked an intensive development of fields that rely on 'omics' data. However, the development in the area of kinetic modeling of large metabolic networks does not follow the same pace. Although metabolomic and fluxomic data are widely available, large-scale kinetic models are very scarce in the literature (Goodacre et al., 2004; Bre-
The efforts in this line of research have been hindered by the lack of pertinent information of the kinetic properties of enzymes in a metabolic network. Experimentally measured kinetic properties of enzymes come necessarily with an extend of uncertainty that originates from differences in experimental conditions and measurement techniques (Wang et al., 2004). In addition, it has been argued that parameter values estimated from the experimental data and directly used into a model are likely to lead to thermodynamical inconsistencies (Liebermeister and Klipp, 2006). Another stumbling block stems from the difficulty to establish rate laws and parameter values for each reaction in large-scale metabolic networks that might contain several hundreds or even thousands of reactions (Jamshidi and Palsson, 2010).

In an attempt to characterize the kinetic responses of metabolic networks in the presence of uncertainty a number of methods that explore the parameter space emerged in the literature (Thomas and Fell, 1994; Petkov and Maranas, 1997; Alves and Savageau, 2000; Almaas et al., 2004). More recently, within the context of Metabolic Control Analysis (MCA), Hatzimanikatis and co-workers have proposed a Monte Carlo sampling procedure for the generation of populations of kinetic models that allow the identification of the rate-limiting steps in metabolic networks (Wang et al., 2004; Wang and Hatzimanikatis, 2006a,b). Motivated by ideas in this work, Liao and co-workers used a Monte Carlo algorithm to sample the parameter space in order to obtain a set of models that all reach the same prescribed steady state (Tran et al., 2008; Rizk and Liao, 2009; Contador et al., 2009). However, this method is computationally intensive and it can be prohibitive for the analysis of large-scale metabolic networks.

Although all aforementioned methods that employ parameter space search circumvent the problem of acquiring kinetic data, there are challenges still to be addressed. The kinetic parameters in each of the reactions are related through the equilibrium constants that are dependent on the standard free energy change of the reaction (Cornish-Bowden and Cardenas, 2000; Nelson and Cox, 2005; Beard and Qian, 2008; Ross, 2008). These dependencies imply that the kinetic parameters’ space is constrained in a very intricate way, which might reduce the sampling efficiency especially in the case of modeling of large-scale metabolic networks. On the other hand, ignoring possible constraints might result in a population of computed models containing a subset of thermodynamically and physico-chemically inconsistent models (Steuer et al., 2006; Grimbs et al., 2007).

In their approach, Hatzimanikatis and co-workers randomly sampled the degree of saturation of the active site of an enzyme and used it subsequently to compute the sensitivities of reaction
responses to the variations of metabolite concentrations and parameters, known within MCA as “elasticities” (Wang et al., 2004; Wang and Hatzimanikatis, 2006a,b). These works laid a foundation for a computational framework for the Optimization and Risk Analysis of Complex Living Entities (ORACLE) that integrates biological information from different sources into a mathematical formalism enabling to identify the rate-limiting steps in biochemical pathways (Mišković and Hatzimanikatis, 2010). In this contribution, we focus on deriving the elasticities as a function of a fractional distribution of the enzyme among its states, the free enzyme and its intermediary complexes, and the displacement from the thermodynamic equilibrium of the individual enzymatic steps. We then exploit the assumption that the total amount of enzyme is preserved in the course of reaction, and we sample randomly the space of enzyme states. Consequently, we use so obtained population of enzyme states, along with the equilibrium displacements of the individual elementary enzymatic steps of the reaction, to compute the elasticities. In our computation, we consider explicitly that the individual equilibrium displacements are constrained by the equilibrium displacement of the overall reaction which, in turn, depends on the thermodynamically feasible metabolite concentrations in the network. So, the proposed computational method explicitly integrates the conservation of the total amount of the enzymes, and the information about reactions thermodynamics and concentration of metabolites.

In contrast to widely used in the literature Markov Chain Monte Carlo methods, the method we propose here does not suffer from the slow mixing property (Schellenberger and Palsson, 2009). This sampling mechanism exploits the specific structure of the underlying constraints and is efficient in the sense that it generates samples of kinetic data that satisfy the imposed constraints and any prescribed distribution of enzyme states at each iteration. The scalability of the sampling mechanism provides means to build genome-scale kinetic models.

**Results and Discussion**

**Fundamental enzyme mechanism**

We study here the reversible Michaelis-Menten enzymatic mechanism as shown in Fig. 1. The reversible Michaelis-Menten kinetics is a three-step mechanism containing three separate steps: (i) binding the enzyme $E$ to the substrate $S$, (ii) the reversible catalytic conversion between the enzyme-substrate $ES$ and the enzyme-product $EP$ complexes, and (iii) binding of the enzyme $E$ to product $P$ (Heinrich and Schuster, 1996). The concentration of the total enzyme $E_T$, the
first-order rate constants $k_{1b}$, $k_{2f}$, $k_{2b}$ and $k_{3f}$, and the second-order rate constants $k_{1f}$ and $k_{3b}$, along with the concentrations of the substrate $S$ and the product $P$, constitute the vector of parameters for this mechanism as follows

$$p = [E_T, k_{1f}, k_{1b}, k_{2f}, k_{2b}, k_{3f}, k_{3b}, S, P]^T.$$  \(1\)

The operator $\cdot^T$ denotes the transpose. In general, $p$ contains three groups of parameters: (i) $p_t$ - conserved concentrations of total enzymes; for single-enzyme reactions it consists of $E_T$, while in the case of substrate channeling it includes the total levels of all involved enzymes; (ii) $p_r$ - rate constants of the elementary reaction steps; and (iii) $p_n$ - other parameters, such as $S$ and $P$ for the reversible Michaelis-Menten mechanism. Therefore, in the more general case the parameter vector, $p$, can be written as:

$$p^T = [p_t^T, p_r^T, p_n^T].$$ \(2\)

It is assumed that the amount of the total enzyme is conserved in the course of reaction

$$E_T = E + ES + EP.$$ \(3\)

For a fixed amount of total enzyme, $E_T$, only two of three enzyme states, $E$, $ES$ and $EP$, can be considered as independent. Assuming that the complexes $ES$ and $EP$ are the independent variables, we derive analytical expressions according to the Metabolic Control Analysis formalism for the reduced stoichiometric matrix $M_R$, the steady-state flux matrix $U$, the elasticity matrices $\Xi$ and $\Pi$ which represent the normalized sensitivities of the individual enzymatic forward and backward reaction rates, $u_{if}$ and $u_{ib}$, $i = 1, 2, 3$ with respect to the concentrations of the intermediate complexes, $ES$ and $EP$, and the parameters, $p$, respectively (see Table I for definitions and Table II for matrix expressions). Subsequently, we assume that the net rate of the reaction is positive, i.e. the reaction operates in the direction from the substrate $S$ to the product $P$, which in turn implies that $\gamma_i \leq 1$, $i = 1, 2, 3$.

From a systems theory perspective, if we consider the single enzyme as the system, and the substrates and products as external inputs (parameters), the derived MCA elements correspond to the concentration and the flux control coefficients with respect to the parameters (including $S$ and $P$), i.e., to the normalized sensitivities of the net reaction rate, $u_{net} = u_{if} - u_{ib}$, with respect to the parameters (cf Table II and Eqs. 11 and 12 in Methods). However, if we consider the enzymes as components of a metabolic network, with $S$ and $P$ as two of its dependent metabolites, i.e.
variables, these control coefficients are in fact the enzyme elasticities with respect to the substrate concentration $S$, the product concentration $P$ and the enzyme activity. In what follows, we will refer to this quantities as metabolite and flux elasticities in order to denote their role within metabolic networks.

**Generating populations of elasticities**

By inspecting Table II, it is observed that if we know the ratios $E/E_T$, $ES/E_T$ and $EP/E_T$, we can compute $\Xi$ and $\Pi$. In general, these ratios are not known. Even when some experimentally observed data are available, the uncertainty introduced due to different experimental conditions makes impossible a precise quantification of these ratios. To generate these quantities, while taking into account the above mentioned uncertainties, we employ the Monte Carlo sampling technique as follows.

Equation 3 divided by $E_T$ defines a simplex in the three-dimensional space of enzyme species as illustrated in Fig. 2. If we assume that the enzyme appears with the equal probability in its three states $E$, $ES$ and $EP$, we can sample uniformly and efficiently the enzyme species over this surface, and thus we can explore the complete three-dimensional subspace of enzyme states that satisfies the constraint in Eq. 3. In the case where the experimental data indicate that the enzyme stays predominantly in one or two of its species, or that any of the enzyme sates follow an observed or hypothesized distribution, we construct an non-uniform distribution over the simplex from which the triplets of the enzyme species are sampled. The resulting random triplets of the ratios $E/E_T$, $ES/E_T$ and $EP/E_T$ can then be used to compute the populations of the matrices $\Xi$ and $\Pi$. The relationships between the enzyme states and the elasticities with respect to the unidirectional rate constants can also be derived as it was done previously in (Kholodenko et al., 1994).

The metabolite and the flux elasticities depend on the equilibrium coefficients $\gamma_i$ and the steady-state net flux $U_{net}$, in addition to the ratios of the enzyme states (see Eqs. 11 and 12 in Methods). The values of the net steady-state flux $U_{net}$ can be estimated using methods from the Flux Balance Analysis (FBA) and Metabolic Flux Analysis (MFA) (Varma and Palsson, 1993a,b; Teusink et al., 2000). On the other hand, the equilibrium displacement $\Gamma$, that can easily be extracted from the experimental data, constrains the equilibrium coefficients $\gamma_i$ (see Methods). This allows to estimate $\gamma_i$'s using the knowledge of $\Gamma$ and the available genomic and kinetic information.
Overall, we show that provided that the equilibrium coefficients and the net steady-state flux are known, the randomly generated samples of the enzyme species triplets allow for computing the populations of flux and metabolite elasticities that correspond to the thermodynamically and physio-chemically feasible states of a reaction. The statistical characteristics of the resulting elasticities can further be analyzed using various data-mining methods.

**Irreversible Michaelis-Menten kinetic mechanism**

The irreversible Michaelis-Menten kinetics has been used in almost every mathematical model of metabolic networks. In this section, we identify the conditions under which the general three-step kinetic mechanism presented in the previous section reduces to the irreversible Michaelis-Menten kinetics.

The irreversible Michaelis-Menten kinetic mechanism characterizes a reaction that kinetically does not favor the binding of the free enzyme to the product. Therefore, for this mechanism the enzyme elasticities with respect to the substrate are uniformly distributed between 0 and 1. In addition, the elasticities with respect to the product are near zero. We investigate under which conditions it is possible to simulate this behavior by using the reaction structure shown in Fig. 1.

In the literature, the common assumption that leads to the use of the irreversible Michaelis-Menten kinetics is that it well describes the kinetic mechanism of the uni-uni reactions away from the equilibrium, i.e. having the equilibrium displacement $\Gamma \ll 1$. However, we find that for a reaction to follow this kinetic mechanism there are additional conditions to be satisfied. In a reaction including three enzymatic reaction steps (Fig. 1), the overall equilibrium displacement is related to the equilibrium coefficients of the individual reaction steps according to the following equation:

$$\Gamma = \gamma_1 \gamma_2 \gamma_3$$  \hspace{1cm} (4)

which implies that the equilibrium coefficients are bounded as follows: $\Gamma \leq \gamma_i \leq 1$ (Fig. 3). We first assume that: (i) the enzyme operates away from the equilibrium, that is $\Gamma = 0.01$, with the equilibrium coefficient for the first enzymatic step being $\gamma_1 \approx \Gamma$, and $\gamma_2 = \gamma_3 \approx 1$; and (ii) the formation of the enzyme-product complex $EP$ is not favorable, i.e. $EP$ appears with a low probability during the course of the reaction.

Based on these two assumptions we generated 2000 random sets of the enzyme states and we analyzed the distributions of the computed populations of elasticities (Fig. 4, panel A). Although this reaction is away from the equilibrium, the resulting elasticities do not correspond to the
irreversible Michaelis-Menten kinetics. The influence of the product concentration is not negligible, and the population of the corresponding flux elasticities is distributed between -1 and 0.

We can better understand the origins of this behavior by deriving the analytical expressions for the flux elasticities:

\[ \varepsilon_{S}^{\text{net}} = \frac{E}{E_T} + \gamma_2 \gamma_3 \frac{E S}{E_T} + \gamma_3 \frac{E P}{E_T} \]
\[ 1 - \gamma_1 \gamma_2 \gamma_3 \]
\[ \varepsilon_{P}^{\text{net}} = - \frac{\gamma_1 \gamma_2 \gamma_3 \frac{E}{E_T} + \gamma_2 \gamma_3 \frac{E S}{E_T} + \gamma_3 \frac{E P}{E_T}}{1 - \gamma_1 \gamma_2 \gamma_3}. \]

(5)

(6)

For reactions away from the equilibrium \((\gamma_1 \gamma_2 \gamma_3 << 1)\), these two expressions reduce to

\[ \varepsilon_{S}^{\text{net}} \approx \frac{E}{E_T} + \gamma_2 \gamma_3 \frac{E S}{E_T} + \gamma_3 \frac{E P}{E_T} \]

\[ \varepsilon_{P}^{\text{net}} \approx - \gamma_2 \gamma_3 \frac{E S}{E_T} - \gamma_3 \frac{E P}{E_T}. \]

These equations lead to two important observations: (i) the elasticities of the enzymes are explicit functions of the distribution of the enzymes among their different states and the individual equilibrium displacements; (ii) for enzymes that operate far from their equilibrium, when the displacement from the equilibrium is primarily contained in the first step \((\gamma_1 \approx \Gamma)\), taking into consideration the conservation of the total enzyme it follows that \(\varepsilon_{S}^{\text{net}} \approx 1\) and \(\varepsilon_{P}^{\text{net}} \approx -\frac{E S}{E_T} - \frac{E P}{E_T}\), which in turn implies that \(\varepsilon_{P}^{\text{net}}\) is distributed between -1 and 0. These observations suggest that large displacement from equilibrium and even high negative Gibbs free energy are not sufficient conditions to describe the reactions following the irreversible Michaelis-Menten kinetics.

From the above analytical expressions for the flux elasticities, one can distinguish several possible scenarios for the validity of irreversible Michaelis-Menten mechanism:

(i) \textit{The last elementary step is far away from the equilibrium, }\gamma_3 << 1, \textit{ whereas the equilibrium coefficients of the first, }\gamma_1, \textit{ and of the intermediary step, }\gamma_2, \textit{ are close to 1 (due to the constraint given in Eq. 4). In addition, the intermediate complex }EP \textit{ appears with a low probability, which in turn results in the equal probability for the enzyme species }E \textit{ and }ES \textit{ to form and dissolve.}

The Gibbs free energy reaction coordinate profiles corresponding to this case are given in Fig. 5. Since the concentration of }EP \textit{ is negligible with respect to the concentrations of }E \textit{ and }ES, \textit{ and considering that }\gamma_3 << 1, \textit{ we have }k_{3b} << k_{3f}. \textit{ Similarly, for the second elementary step where the equilibrium coefficient }\gamma_2 \approx 1 \textit{ the conversion of }EP \textit{ to }ES \textit{ is kinetically more favorable than the conversion in the opposite direction, i.e. }k_{2f} << k_{2b}. \textit{ These cases are}
illustrated in the right-hand part of the energy diagram (Fig. 5). We can see that the free activation energies of dissociation of the complex \( EP \), which are inversely proportional to the kinetic rate constants \( k_{2b} \) and \( k_{3f} \), are small. In turn, this leads to infinitesimal levels of concentration of the enzyme-product complex. In contrast to the intermediate and the last elementary step where the ratio between kinetic rate constants are well defined, the ratio of the kinetic constants in the first step can vary leading to several alternative energy profiles for \( ES \) as shown in Fig. 5. Depending on the concentration levels of \( E \) and \( ES \) the first elementary step can be kinetically more favorable in the forward or in the backward direction, or to be in the kinetic equilibrium. This scenario corresponds to the structure of irreversible Michaelis-Menten, well-known in the literature (Heinrich and Schuster, 1996), where an enzyme appears predominantly in the reversibly interconvertible forms of the free enzyme \( E \) and the enzyme-substrate complex \( ES \), and the product \( P \) is irreversibly produced from the enzyme-substrate complex.

(ii) The last elementary step has large displacement from its equilibrium \( (\gamma_3 \ll 1) \), \( \gamma_1 = \gamma_2 \approx 1 \), and the ES intermediate complex appears with a low probability. Under these conditions we have \( k_{1f} << k_{1b} \) and \( k_{2f} >> k_{2b} \), which implies that the enzyme-substrate complex, \( ES \), dissociates rapidly and that its concentration levels are very low. The corresponding energy diagrams are represented in the panel A of Fig. 6. In the last elementary step, there are several alternative activation energy profiles depending on the relative concentration of the free enzyme with respect to the enzyme-product complex, \( EP \).

(iii) The intermediary step has large displacement from the equilibrium \( (\gamma_2 \ll 1) \), \( \gamma_1 \) and \( \gamma_3 \) are close to 1, and the probability of the formation of \( EP \) is low. The Gibbs energy profiles depicting this type of reactions are given in the panel B of Fig. 6. These energy profiles are similar to the ones shown in Fig. 5 with a difference in the ratios between \( k_{3b} \) and \( k_{3f} \), and \( k_{2b} \) and \( k_{2f} \) which are slightly less pronounced. Consequently, the free activation energies of the dissociation of \( EP \) are slightly increased.

We have generated populations of the enzyme species and computed the elasticities according to the conditions of three scenarios (Table III). In all three cases, the distributions of the elasticities match the irreversible Michaelis-Menten kinetics (Fig. 4, panel B: distribution for scenario (i); distributions for scenarios (ii) and (iii) are similar and not shown). We observe, as expected, that the flux elasticities with respect to the substrate are uniformly distributed between 0 and 1, and that the ones with respect to the product are close to zero.
Using partial information

We have shown that when no a priori information about a reaction is available it is possible to generate populations of corresponding elasticities by sampling efficiently a simplex in the space of enzyme states. Any additional experimental information about the reaction will help in reducing the sampling space which in turn will allow more accurate computational predictions of the responses of the enzymes to changes in metabolites concentrations and parameters. This exchange of information between computational and experimental studies will result in both better computational models and better design of experiments. Toward this end, we developed a framework to incorporate the additional experimental data in the procedure of computing the elasticities.

As an illustration, we consider again the reaction consisting of three enzymatic elementary steps (Fig. 1). We assume that experimentally observed data suggest that the enzyme appears in its free form $E$ with the probability of 60%, and in the form of enzyme-metabolites complexes $ES$ and $EP$ with the probabilities of 15% and 25%, respectively. We approximate the observed distribution of the experimental data with the Dirichlet distribution whose parameters can be estimated based on the mean values and variances of the acquired data. When estimating the parameters of Dirichlet distribution we consider the fact that the bigger the sum of the parameters is, the smaller the variance becomes. Consequently, by multiplying the vector of parameters with a positive constant bigger than 1, one can reduce the variance of the distribution. More details of this procedure are given in Methods and references therein. The probability density function $f_D$ that approximates the experimental distribution is shown in Fig. 7, panel A, and the corresponding 2000 generated samples are shown in panel B. It is observed that the prescribed marginal distributions for $E$, $ES$ and $EP$ are well approximated. In Fig. 7, panel B, we also show 2000 data points generated through uniform sampling of simplex which can be used in the case when experimental information is not available. As expected, the density of sampling when partial information is included is higher which allows more thorough characterization of the kinetic space.

Conclusions

In spite of advances in measurement techniques, development of kinetic models is a challenging task because kinetic data are still of limited availability and come with the inevitable uncertainty. The variability and incompleteness of data call upon optimization and risk analysis methods such as ORACLE (Mišković and Hatzimanikatis, 2010) that are able to quantify the uncertainty and
to integrate available experimental data into models providing reliable predictions/expectations of the behavior of biochemical pathways in the face of perturbations. The formalism presented here allows us to generate and study all biochemically and thermodynamically plausible kinetic models of biochemical reactions thus paving the way for the development of the framework for uncertainty modeling of the kinetics in biochemical pathways. Specific mathematical formulation of the underlying problem allows for efficient generation of populations of kinetic models and enables to scale this methodology even to genome-scale biochemical pathways.

Methods

Systems-oriented models of enzyme kinetics

Assuming a spatial homogeneity in a biochemical reaction, the dynamic behavior of the concentrations of enzyme states, such as $E$, $ES$ and $EP$, can be described by a set of ordinary differential equations of the following form:

$$\frac{dx_e}{dt} = Mu(x_e, p), \quad (9)$$

where $x_e$ represents the $m$-dimensional vector of the concentrations of enzyme species, $M$ denotes the $m \times n$-dimensional matrix describing the stoichiometry of the reaction, $u$ is the $n$-dimensional vector of metabolic fluxes within this reaction, referred in the sequel as the internal metabolic fluxes, and $p$ is the vector of parameters defined in Eq. 2. In this mathematical representation, a reversible flux is expressed as a difference between the corresponding forward irreversible and backward irreversible fluxes (Wang et al., 2004).

The conservation relation between enzyme states introduces a rank deficiency of the stoichiometric matrix $M$. To overcome this, the vector of concentrations $x_e$ can be decomposed in two sub-vectors: an independent enzyme states concentration vector $x_{ei}$, and a dependent enzyme states concentration vector $x_{ed}$ (Reder, 1988). In addition, the rows in the stoichiometric matrix $M$ corresponding to the mass balances of the dependent enzyme states can be eliminated so that a new, reduced, stoichiometric matrix $M_R$ is formed (Heinrich and Schuster, 1996; Reder, 1988). Therefore, the mass balance equations of the metabolic reaction (Eq. 9) are reduced to the following form

$$\frac{dx_{ei}}{dt} = M_R u(x_{ei}, x_{ed}(x_{ei}, p), p). \quad (10)$$
Elasticities

We apply the log(linear) kinetic formalism of MCA for the calculation of elasticities in a (bio)chemical reaction (Hatzimanikatis et al., 1996; Hatzimanikatis and Bailey, 1996, 1997). Metabolite elasticities $\varepsilon^x_p$, and flux elasticities $\varepsilon^u_p$ quantify the variations of enzyme states concentrations and internal metabolic fluxes, respectively, with respect to the variations in system parameters. As discussed earlier, if one considers the single enzyme as the system, with the concentrations of the metabolites, $S$ and $P$, being the parameters, these quantities are in fact the control coefficients. From the same perspective, the enzyme states can be considered as metabolites. Nevertheless, we consider the enzymes as a part of metabolic networks, hence we refer $\varepsilon^x_p$ and $\varepsilon^u_p$ as the elasticities. The expressions for these elasticities can be derived by linearizing and scaling the system (Eq. 10) around the steady-state:

$$\varepsilon^x_{ei} = - (MRU\Xi)^{-1} MRU\Pi$$
$$\varepsilon^u_p = \Xi \varepsilon^{x_{ei}}_{p} + \Pi$$

where $U$ denotes the diagonal matrix of the internal steady-state fluxes, and $\Xi = \Xi_i + \Xi_d Q^d_i$ is the matrix of the local sensitivities with respect to the enzyme state concentrations; $\Xi_i$ and $\Xi_d$ are the matrices of the elasticities with respect to independent and dependent enzyme states, respectively, and $Q^d_i$ represents the relative abundance of the dependent enzyme states compared to that of the independent ones. The matrix $\Pi = [\Pi_t, \Pi_r, \Pi_n]$ denotes the local sensitivities with respect to the system parameters, $p^T = [p^T_t, p^T_r, p^T_n]$. The matrix $\Pi_t$ can be expressed as $\Pi_t = \Xi_d Q^d_t$ where the weight matrix $Q^d_t$ represents the relative abundance of dependent enzyme states with respect to the amounts of their corresponding total enzymes.

Equilibrium factors

We define the equilibrium coefficient as the ratio of the backward and the forward reaction rates $u_b$ and $u_f$, respectively:

$$\gamma = \frac{u_b}{u_f} = \frac{u_f - u_{net}}{u_f}$$

with $u_{net}$ denoting the net flux rate. The equilibrium coefficient, $\gamma \in [0, +\infty)$, reflects the reversibility of a reaction with respect to the net flux. Values of $\gamma$ close to zero, $\gamma \to 0$, or close to infinity, $\gamma \to \infty$, indicate forward irreversible and backward irreversible reaction steps, respectively. In contrast, the values of $\gamma$ close to 1, $\gamma \approx 1$, imply that the net flux is negligible with respect
to backward and forward fluxes, i.e. the enzyme operates near thermodynamic equilibrium.

For a reaction with one product and one substrate, the displacement from the thermodynamic equilibrium can be defined as follows:

\[
\Gamma = \frac{S_{eq} P}{P_{eq} S} = \frac{1}{k_{eq}} \frac{P}{S}
\]

(14)

with \(S_{eq}\) and \(P_{eq}\) being respectively the concentration of the substrate and the product at the thermodynamic equilibrium (Heinrich and Schuster, 1996). \(k_{eq}\) denotes the equilibrium constant defined as a ratio between the products of the forward and the backward rate constants, respectively. Without any loss of generality, we assume that there is a net production of the product \(P\), i.e. \(\Gamma < 1\) or \(P/S < k_{eq}\), which implies that the Gibbs free energy difference is negative, i.e. \(\Delta G < 0\). The displacement from the thermodynamic equilibrium \(\Gamma\) can readily be extracted from the experimentally observed data.

Monte Carlo Sampling Over the Constrained Kinetic Space

We have shown that by sampling the enzyme states while preserving the conservation of the amount of the total enzyme, the uncertainty in the elasticities can be modeled and analyzed statistically in order to characterize the rate-limiting steps. We further extend the proposed idea to address the more general case, to a more complex group of kinetic mechanisms where the enzyme in its free form and enzyme complexes appear linearly in the rate equations. For this kind of kinetic mechanisms, the matrices \(\Xi\) and \(\Pi\), as in the case of the irreversible Michaelis-Menten kinetics, can be expressed as functions of ratios of the enzymes states and the amount of the total enzyme \(E_T\). Hence, following the reasoning from the previous section, we can establish a general Monte Carlo methodology allowing us to quantify the elasticities. We assume that details about the kinetic mechanism are available, or, an appropriate kinetics is assumed. The procedure is as follows:

1. Construct the stoichiometry matrix using available biochemical and genomics information, and separate the enzyme states into an independent and a dependent group.

2. Estimate steady-state fluxes using FBA methods.

3. Generate the random enzyme states samples satisfying the conservation constraints on the amount of total enzyme. If no a priori information about the distribution of the enzyme states is available, the random samples are uniformly generated. Information from acquired
experimental data can be included to introduce either a bias in the uniform distribution or even to define other types of distribution.

4. Determine the equilibrium coefficients on the basis of the nature of kinetic mechanism and the knowledge of the equilibrium displacement $\Gamma$. If only $\Gamma$ is known, the bounds on the equilibrium coefficients are precisely defined, thus the space of the equilibrium coefficients can be sampled to generate the populations of the thermodynamically feasible realizations of the underlying reaction.

5. Construct the sensitivity matrices $\Xi, \Xi_d$ and $\Pi$ followed by the computation of the metabolite and flux elasticities according to Eqs. 11 and 12.

In the sequel, we give more details about the generation of the random enzyme states samples mentioned in the third step of the procedure.

**Uniform sampling of enzyme states**

There exists an abundant literature on random variate generation over different regions (Feller, 1968; Rubinstein, 1981; Devroye, 1986; Garvey, 2000; Gentle, 2003). The most common methods used for uniform sampling are acceptance-rejection methods (Von Neumann, 1963), and Monte Carlo Markov Chain methods (Gilks et al., 1998; Brémaud, 1999). However, the former methods tend to be very inefficient in high dimension space, while the latter require a large number of samples to obtain asymptotically uniform coverage of the space.

In this paper we make use of ideas presented in (Wilks, 1962; Rubinstein, 1982) that exploit the particular structure of the enzyme space we want to sample uniformly. For a general kinetic mechanism, where an enzyme appears in the free form $E$ and in the form of $n-1$ different enzyme-metabolite complexes $EX_i$, $i = 1, \ldots, n-1$, we can write

$$E + \sum_{i=1}^{n-1} EX_i = ET \Rightarrow \frac{E}{ET} + \sum_{i=1}^{n-1} \frac{EX_i}{ET} = 1.$$  \hspace{1cm} (15)

In the $n$-dimensional space of ratios $(E/ET, EX_1/ET, EX_2/ET, \ldots, EX_{n-1}/ET)$, the constraint given in Eq. 15 represents a simplex. Observe also that $E/ET$ and $EX_i/ET$, $i = 1, \ldots, n-1$ are constrained between 0 and 1. An example of a simplex in the three-dimensional space is shown in Fig. 2.

We next use the fact that a random vector uniformly distributed over $n$-dimensional simplex can be obtained by generating samples from a $n$-variate Dirichlet distribution with all parameters
equal to 1, which will be referred subsequently as $D(1, \ldots, 1)$. Generation of random variables from $D(1, \ldots, 1)$ can then be performed as follows (Rubinstein, 1982).

**Algorithm 1** (*Generating uniform enzyme states*)

1. Generate a random vector $(X_1, \ldots, X_n)$ exponentially distributed with the rate parameter equal to 1.

2. Compute the random vector

$$(Y_1, \ldots, Y_n) = \left( \frac{X_1}{\sum_{i=1}^{n} X_i}, \ldots, \frac{X_n}{\sum_{i=1}^{n} X_i} \right).$$

The random vector $(Y_1, \ldots, Y_n)$ is distributed with $D(1, \ldots, 1)$, i.e. it is uniformly distributed on the surface of $n$-dimensional simplex. So, Algorithm 1 provides a very efficient and simple way to generate sets of enzyme states $(E/E_T, EX_1/E_T, EX_2/E_T, \ldots, EX_{n-1}/E_T)$ even in the case of high-dimensional spaces.

**Non-uniform sampling of enzyme states**

Detailed knowledge of ranges of probability with which an enzyme appears in each of its states can be used to generate more refined sets of enzyme states. Once again, we benefit from the fact that Dirichlet distribution samples naturally from simplex. In contrast to the case of uniform sampling where all parameters of Dirichlet distribution were equal to 1, in this case we change these parameters so that the shape of Dirichlet distribution approximatively corresponds to the experimentally observed data.

The probability density function of the Dirichlet distribution with parameters $\nu_1, \nu_2, \ldots, \nu_{n-1}, \nu_n$ is given as

$$f_D(\nu_1, \ldots, \nu_n) = \frac{\Gamma \left( \sum_{i=1}^{n} \nu_i \right)}{\prod_{i=1}^{n} \Gamma(\nu_i)} \prod_{i=1}^{n-1} \sigma_i^{\nu_i-1} \left( 1 - \sum_{i=1}^{n-1} \sigma_i \right)^{\nu_n-1}$$

where $\Gamma(.)$ denotes the gamma function, the random variables $\sigma_1, \ldots, \sigma_{n-1} > 0$ satisfy $\sigma_1 + \ldots + \sigma_{n-1} < 1$, and the parameters $\nu_1, \ldots, \nu_n > 0$. Outside the simplex the probability density function is zero. The Dirichlet distribution has a nice property that can be used in approximating the experimental data. For a random vector $(Y_1, \ldots, Y_n)$ distributed with $D(\nu_1, \ldots, \nu_n)$ we have
that the mean values of each component of the random vector are given as

\[ E \{ Y_i \} = \frac{\nu_i}{\sum_{k=1}^{n} \nu_k} \]

and the corresponding variances read as

\[ \text{Var} \{ Y_i \} = \frac{\nu_i \left( \sum_{k=1, k \neq i}^{n} \nu_k \right)}{(1 + \sum_{k=1}^{n} \nu_k) \left( \sum_{k=1}^{n} \nu_k \right)^2}. \]

In addition, the marginal distributions of \( X_i \)'s are Beta distributions, i.e. \( Y_i \sim \text{Beta}(\nu_i, \nu_0 - \nu_i) \). Hence, we can use information extracted from the experimental data, such as the ranges of the probability in which each of enzyme states appear, to define a mean value and a variance for each of the marginal distributions of enzyme states. So defined mean values and variances are subsequently used to compute the parameters of Dirichlet distribution. In other words, we design the Dirichlet distribution through shaping its marginal distributions. In the case of high-dimensional enzyme states space, the problem of finding the parameters \( \nu_1, \ldots, \nu_n \) can be cast as an optimization problem.

Once having determined the parameters of the Dirichlet distribution that approximates well the experimental data, we turn to the problem of generating the samples of that distribution. In (Arnason, 1972), an efficient method for generating the Dirichlet variates with given parameters \( \nu_1, \nu_2, \ldots, \nu_{n-1}, \nu_n \) is proposed. This method is described by the following algorithm.

**Algorithm 2 (Generating non-uniformly sampled enzyme states)**

1. Generate a random vector \((X_1, \ldots, X_n)\) distributed with the gamma distributions \(G(\nu_i), i = 1, \ldots, n\).

2. Compute the Dirichlet variates as follows:

\[ (Y_1, \ldots, Y_n) = \left( \frac{X_1}{\sum_{i=1}^{n} X_i}, \ldots, \frac{X_n}{\sum_{i=1}^{n} X_i} \right). \]

The Dirichlet variates \((Y_1, \ldots, Y_n)\) have the probability density function \(D(\nu_1, \ldots, \nu_n)\). Algorithm 2 is a general version of Algorithm 1 and can be used to generate uniform variates as well.
References


Tables

Table I. Definitions.

<table>
<thead>
<tr>
<th>Name</th>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metabolite elasticities</td>
<td>$\varepsilon_{ei}^x$</td>
<td>$\frac{dlnx_{ei}}{dlnp}$</td>
</tr>
<tr>
<td>Flux elasticities</td>
<td>$\varepsilon_{ei}^u$</td>
<td>$\frac{dlnu}{dlnp}$</td>
</tr>
<tr>
<td>Independent enzyme states elasticities</td>
<td>$\Xi_i$</td>
<td>$\frac{dlnu}{dlnx_{ei}}$</td>
</tr>
<tr>
<td>Dependent enzyme states elasticities</td>
<td>$\Xi_d$</td>
<td>$\frac{dlnu}{dlnx_{ed}}$</td>
</tr>
<tr>
<td>Total enzyme parameter elasticities</td>
<td>$\Pi_t$</td>
<td>$\frac{dlnu}{dlnp_t}$</td>
</tr>
<tr>
<td>Reaction rates elasticities</td>
<td>$\Pi_r$</td>
<td>$\frac{dlnu}{dlnp_r}$</td>
</tr>
<tr>
<td>Other parameters elasticities</td>
<td>$\Pi_n$</td>
<td>$\frac{dlnu}{dlnp_n}$</td>
</tr>
<tr>
<td>Relative abundance</td>
<td>$Q_i^d$</td>
<td>$\frac{dlnx_{ei}}{dlnx_{ed}}$</td>
</tr>
<tr>
<td>Total enzyme weights</td>
<td>$Q_t$</td>
<td>$\frac{dlnx_{ed}}{dlnp_t}$</td>
</tr>
</tbody>
</table>
Table II. Elasticities’ constitutive elements for the fundamental three-step reversible kinetic mechanism (Fig. 1). In the expression for the steady-state flux matrix $U$, $U_{net}$ denotes the steady-state net flux, while $\gamma_i$, $i = 1, 2, 3$ are the equilibrium coefficients that represent the ratio between the backward reaction rates, $u_{ib}$, and the forward ones, $u_{if}$, in each of the elementary reaction steps (see Methods): $\gamma_1 = u_{1b}/u_{1f}$, $\gamma_2 = u_{2b}/u_{2f}$ and $\gamma_3 = u_{3b}/u_{3f}$.

$$M_R = \begin{pmatrix} 1 & -1 & -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 & -1 & 1 \end{pmatrix} \quad \Xi = \begin{pmatrix} -ES/E & 1 & 1 & 0 & 0 & -ES/E \\ -EP/E & 0 & 0 & 1 & 1 & -EP/E \end{pmatrix}^T$$

$$U = \begin{pmatrix} 1/\gamma_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \gamma_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/\gamma_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & \gamma_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & \gamma_3 & 0 \\ 0 & 0 & 0 & 0 & 0 & \gamma_3 \end{pmatrix} \quad U_{net} \quad \Pi = \begin{pmatrix} E_T/E & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & E_T/E & 0 & 1 \end{pmatrix} \quad I_{6\times6}$$

Table III. Conditions under which the three-step kinetic mechanism reduces to the irreversible Michaelis-Menten kinetics.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>$\gamma_1$</th>
<th>$\gamma_2$</th>
<th>$\gamma_3$</th>
<th>$E$</th>
<th>$ES$</th>
<th>$EP$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i)</td>
<td>$\approx 1$</td>
<td>$\approx 1$</td>
<td>$&lt;&lt; 1$</td>
<td>uniformly distributed</td>
<td>uniformly distributed</td>
<td>$\approx 0$</td>
</tr>
<tr>
<td>(ii)</td>
<td>$\approx 1$</td>
<td>$\approx 1$</td>
<td>$&lt;&lt; 1$</td>
<td>uniformly distributed</td>
<td>$\approx 0$</td>
<td>uniformly distributed</td>
</tr>
<tr>
<td>(iii)</td>
<td>$\approx 1$</td>
<td>$&lt;&lt; 1$</td>
<td>$\approx 1$</td>
<td>uniformly distributed</td>
<td>uniformly distributed</td>
<td>$\approx 0$</td>
</tr>
</tbody>
</table>
Figure Legends

Fig. 1 Reversible Michaelis-Menten kinetics consisting of three enzymatic reaction steps.

Fig. 2 Example of a three-dimensional enzyme states space. The enzyme appears in the free form $E$, as the enzyme-substrate complex $ES$ and as the enzyme-product complex $EP$. The enzyme species are randomly generated over the simplex.

Fig. 3 The surface $\Gamma = \gamma_1\gamma_2\gamma_3$ in the space of elementary displacements $\gamma_1$, $\gamma_2$ and $\gamma_3$. As one approaches to the thermodynamic equilibrium the surface of $\Gamma$ shrinks, and as a consequence the allowable ranges of $\gamma_1$, $\gamma_2$ and $\gamma_3$ shrink as well.

Fig. 4 Histograms of the flux elasticities for the three-step reaction away from the equilibrium ($\Gamma = 0.01$). Panel A: the flux elasticities with respect to the substrate $S$ (upper part) and the product $P$ (lower part) concentrations; the equilibrium coefficients are $\gamma_1 = 0.0104$, $\gamma_2 = 0.98$, and $\gamma_3 = 0.98$; Panel B: histograms of elasticities for the irreversible Michaelis-Menten kinetics - the equilibrium coefficients are $\gamma_1 = \gamma_2 = 0.98$, and $\gamma_3 = 0.0104$.

Fig. 5 Gibbs free energy profiles for the reaction following the irreversible Michaelis-Menten kinetics. The last elementary step is largely displaced from the equilibrium ($\gamma_3 << 1$), while the other two steps are close to the equilibrium. High values of the rate constants $k_3f$ and $k_2b$ result in a very low concentration levels of $EP$. The Gibbs energy levels of substrate, intermediates and product are given in dark blue, whereas the transition states are traced in red. For different pairs of concentration levels for $E$ and $ES$ several alternative energy configurations can be distinguished as shown with light gray curves. The corresponding energy levels are traced in light blue.

Fig. 6 Reaction energy profiles for the reactions characterized in scenarios (ii) and (iii). Panel A: High values for $k_{1b}$ and $k_2f$ constants drive the concentration levels of $ES$ very low. The equilibrium displacements $\gamma_3 << 1$, $\gamma_1 = \gamma_2 \approx 1$. Panel B: The equilibrium displacement predominantly expressed in the intermediary reaction step. The concentration levels of $EP$ substantially smaller than the ones of $E$ and $ES$. The alternative energy levels for $ES$ shown in light blue.

Fig. 7 Non-uniform sampling of enzyme space. Panel A: probability density function $f_D$ as a function of the simplex $\frac{E}{E_T} + \frac{ES}{E_T} + \frac{EP}{E_T} = 1$ lying in the horizontal plane. This distribution
allows for generating random sets of enzyme states for a reaction where the enzyme is in its forms $E$, $ES$ and $EP$ with probability of 60%, 15% and 25%, respectively: Panel B: uniform sampling (dots) and biased sampling (diamonds) with the probability density function $f_D$ presented in Panel A. Samples presented in the space ($E/E_T$, $ES/E_T$, $EP/E_T$).
Figures

Figure 1. Reversible Michaelis-Menten kinetics consisting of three enzymatic reaction steps.

Figure 2. Example of a three-dimensional enzyme states space. The enzyme appears in the free form $E$, as the enzyme-substrate complex $ES$ and as the enzyme-product complex $EP$. The enzyme species are randomly generated over the simplex.
Figure 3. The surface $\Gamma = \gamma_1 \gamma_2 \gamma_3$ in the space of elementary displacements $\gamma_1$, $\gamma_2$ and $\gamma_3$. As one approaches to the thermodynamic equilibrium the surface of $\Gamma$ shrinks, and as a consequence the allowable ranges of $\gamma_1$, $\gamma_2$ and $\gamma_3$ shrink as well.

Figure 4. Histograms of the flux elasticities for the three-step reaction away from the equilibrium ($\Gamma = 0.01$). Panel A: the flux elasticities with respect to the substrate $S$ (upper part) and the product $P$ (lower part) concentrations; the equilibrium coefficients are $\gamma_1 = 0.0104$, $\gamma_2 = 0.98$, and $\gamma_3 = 0.98$; Panel B: histograms of elasticities for the irreversible Michaelis-Menten kinetics - the equilibrium coefficients are $\gamma_1 = \gamma_2 = 0.98$, and $\gamma_3 = 0.0104$. 
Figure 5. Gibbs free energy profiles for the reaction following the irreversible Michaelis-Menten kinetics. The last elementary step is largely displaced from the equilibrium ($\gamma_3 << 1$), while the other two steps are close to the equilibrium. High values of the rate constants $k_{3f}$ and $k_{2b}$ result in a very low concentration levels of $EP$. The Gibbs energy levels of substrate, intermediates and product are given in dark blue, whereas the transition states are traced in red. For different pairs of concentration levels for $E$ and $ES$ several alternative energy configurations can be distinguished as shown with light gray curves. The corresponding energy levels are traced in light blue.
Figure 6. Reaction energy profiles for the reactions characterized in scenarios (ii) and (iii). Panel A: High values for $k_{1b}$ and $k_{2f}$ constants drive the concentration levels of ES very low. The equilibrium displacements $\gamma_3 << 1$, $\gamma_1 = \gamma_2 \approx 1$. Panel B: The equilibrium displacement predominantly expressed in the intermediary reaction step. The concentration levels of EP substantially smaller than the ones of E and ES. The alternative energy levels for ES shown in light blue.

Figure 7. Non-uniform sampling of enzyme space. Panel A: probability density function $f_D$ as a function of the simplex $\frac{E}{E_T} + \frac{ES}{E_T} + \frac{EP}{E_T} = 1$ lying in the horizontal plane. This distribution allows for generating random sets of enzyme states for a reaction where the enzyme is in its forms E, ES and EP with probability of 60%, 15% and 25%, respectively; Panel B: uniform sampling (dots) and biased sampling (diamonds) with the probability density function $f_D$ presented in Panel A. Samples presented in the space ($E/E_T$, $ES/E_T$, $EP/E_T$).